Econometrica Supplementary Material

SUPPLEMENT TO "HERDING AND CONTRARIAN BEHAVIOR IN FINANCIAL MARKETS"

(*Econometrica*, Vol. 79, No. 4, July 2011, 973–1026)

BY ANDREAS PARK AND HAMID SABOURIAN

There are results in the paper that were not fully discussed or fully proven. This Supplementary Material contains material that was omitted or mentioned. We organize this appendix in the same way as the sections in the main paper.

APPENDIX A: SUPPLEMENTARY MATERIAL FOR SECTION 5

A.1. *Proof of Lemma 1*

Observe first that

$$\begin{aligned} &\mathsf{E}[V|S,H^t]-\mathsf{E}[V|H^t] \\ &\quad = \mathcal{V}q_2^t\left(\frac{\Pr(S|V_2)}{\Pr(S)}-1\right)+2\mathcal{V}q_3^t\left(\frac{\Pr(S|V_3)}{\Pr(S)}-1\right). \end{aligned}$$

The right hand side of the above equality has the same sign as

$$\begin{aligned} &q_2^t\left(\Pr(S|V_2)\sum_j q_j^t-\sum_j \Pr(S|V_j)q_j^t\right) \\ &\quad + 2q_3^t\left(\Pr(S|V_3)\sum_j q_j^t-\sum_j \Pr(S|V_j)q_j^t\right) \\ &= q_1^t q_2^t(\Pr(S|V_2)-\Pr(S|V_1))+q_2^t q_3^t(\Pr(S|V_2)-\Pr(S|V_3)) \\ &\quad + 2q_3^t\big(q_1^t\big(\Pr(S|V_3)-\Pr(S|V_1)+q_2^t(\Pr(S|V_3)-\Pr(S|V_2))\big)\big). \end{aligned}$$

Q.E.D.

A.2. *Proof of Lemma 2*

The proof follows by Lemma 1: By the symmetry assumption on the priors $(q_1^1=q_3^1)$, equation (2) is negative (positive) at $t=1$ if and only if $(\Pr(S|V_3)-\Pr(S|V_1))(q_2^1+2q_1^1)q_3^1$ is less (greater) than 0; the latter is equivalent to S having a negative (positive) bias. *Q.E.D.*

A.3. *Proof of Lemma 3*

The claim follows from $\mathsf{E}[V|H^t]-\mathsf{E}[V]=\mathcal{V}[(1-q_1^t-q_3^t)+2q_3^t]-\mathcal{V}=\mathcal{V}(q_3^t-q_1^t)$. *Q.E.D.*

 DOI: 10.3982/ECTA8602

A.4. *Proof of Lemma 9*

The proof is analogous to the derivation in the proof of Lemma 1. To show the lemma note that

$$\mathsf{E}[V|S,H^t]-\mathsf{ask}^t=\mathcal{V}q_2\left(\frac{\mathsf{Pr}(S|V_2)}{\mathsf{Pr}(S)}-\frac{\beta_2}{\mathsf{Pr}(\mathsf{buy}|H^t)}\right)+2\mathcal{V}q_3\left(\frac{\mathsf{Pr}(S|V_3)}{\mathsf{Pr}(S)}-\frac{\beta_3}{\mathsf{Pr}(\mathsf{buy}|H^t)}\right).$$

The right hand side of the above expression has the same sign as

$$\begin{aligned}
&q_2\left(\mathsf{Pr}(S|V_2)\sum_j\beta_jq_j-\beta_2\sum_j\mathsf{Pr}(S|V_j)q_j\right)\\
&\quad+2q_3\left(\mathsf{Pr}(S|V_3)\sum_j\beta_jq_j-\beta_3\sum_j\mathsf{Pr}(S|V_j)q_j\right)\\
&=q_1q_2(\beta_1\mathsf{Pr}(S|V_2)-\beta_2\mathsf{Pr}(S|V_1))\\
&\quad+q_2q_3(\beta_3\mathsf{Pr}(S|V_2)-\beta_2\mathsf{Pr}(S|V_3))\\
&\quad+2q_3(q_1(\beta_1\mathsf{Pr}(S|V_3)-\beta_3\mathsf{Pr}(S|V_1))\\
&\quad+q_2(\beta_2\mathsf{Pr}(S|V_3)-\beta_3\mathsf{Pr}(S|V_2))).
\end{aligned}$$

Q.E.D.

A.5. *Necessity of the Bounds on Noise Trading*

In the paper, we claimed that the bounds on noise trading are also necessary for some cases. The following proposition outlines these scenarios.

PROPOSITION 3a: (i) *Suppose that S buy herds and that there is at most one U-shaped signal. Then $\mu<\min\{\mu^i,\mu^s_{bh}\}$, where μ^i is defined in Lemma* 8 *and μ^s_{bh} is defined in* (12).

(ii) *Suppose that S acts as a buy contrarian and there is at most one hill-shaped signal. Then $\mu<\min\{\mu^i,\mu^s_{bc}\}$, where μ^i is defined in Lemma* 8 *and μ^s_{bc} is defined in* (13).

PROOF: We shall prove (i); the proof of (ii) is analogous. Assume S buy herds. Then S sells initially. It follows from Lemma 5 that $\mu<\mu^i$.

To show that $\mu<\mu^s_{bh}$, first note that by Proposition 2, S must be nU-shaped. Next consider the different possibilities separately.

Case A—There is no signal $S'\neq S$ such that $\mathsf{Pr}(S'|V_3)>\mathsf{Pr}(S'|V_2)$. Then it must be that $\mu_2(S')=1$ for all S' and, therefore, it must be that $\mu<\mu^s_{bh}=1$.

Case B—There is a signal $S'\neq S$ such that $\mathsf{Pr}(S'|V_3)>\mathsf{Pr}(S'|V_2)$. Since S is U-shaped, it must be that $\mathsf{Pr}(S|V_3)>\mathsf{Pr}(S|V_2)$ and $\mathsf{Pr}(S''|V_3)\le\mathsf{Pr}(S''|V_2)$ for $S''\neq S,S'$. This implies that $\mu_2(S'')=1$ and hence $\mu_2(S')=\mu^s_{bh}$.

Now there are two cases. First, if $\mu_2(S')$ also equals 1, then clearly $\mu^s_{bh} = 1$ and the claim is trivially true.

Second, assume that $\mu_2(S') = \mu^s_{bh} < 1$. Since S buy herds at H^t, to show that $\mu < \min\{\mu^i, \mu^s_{bh}\}$, it suffices to show that S' also buys whenever S buys (the alternative is that S' does not buy so that $\mu_2(S') = 1 > \mu^s_{bh}$). When S' buys, $\mathsf{E}[V|S', H^t] - \mathsf{ask}^t > 0$. Suppose S' does not buy. As the sign of $\mathsf{E}[V|S', H^t] - \mathsf{ask}^t$ is given by equation (11), it must then hold that

$$
\begin{aligned}
\text{(S1)} \qquad & q_1 q_2[\beta_1 \mathsf{Pr}(S'|V_2) - \beta_2 \mathsf{Pr}(S'|V_1)] + q_2 q_3[\beta_2 \mathsf{Pr}(S'|V_3) - \beta_3 \mathsf{Pr}(S'|V_2)] \\
& \quad + 2q_1 q_3[\beta_1 \mathsf{Pr}(S'|V_3) - \beta_3 \mathsf{Pr}(S'|V_1)] \le 0.
\end{aligned}
$$

Also, since there is at most one U-shaped signal, it must be that

$$
\text{(S2)} \qquad \mathsf{Pr}(S'|V_3) > \mathsf{Pr}(S'|V_2) \ge \mathsf{Pr}(S'|V_1).
$$

By Proposition 1, this implies that S' does not sell. By supposition, S' does not buy and, therefore, S is the only buyer at H^t (S'' is selling). Since S is nU-shaped, we must also have $\beta^t_1 > \beta^t_3 \ge \beta^t_2$. This, together with (S2), implies that the first and the third terms in (S1) are positive. Furthermore, the second term has the same sign as

$$
\text{(S3)} \qquad \gamma(\mathsf{Pr}(S'|V_3) - \mathsf{Pr}(S'|V_2)) + \mu(\mathsf{Pr}(S|V_2)\,\mathsf{Pr}(S'|V_3) - \mathsf{Pr}(S|V_3)\,\mathsf{Pr}(S'|V_2)).
$$

By (S2), the first term in the last expression is positive; furthermore, since S is nU, we have $m^2 = \mathsf{Pr}(S|V_3) > \mathsf{Pr}(S|V_2)$. Since $\mu_2(S') < 1$, we must have that $M^2(S') < 1$ is negative. But $-\mu M^2(S')$ is the second term in the last expression and it is thus positive. Consequently, (S3) is positive. Therefore, the second term in (S1) must also be positive. Therefore, S' must be buying at any H^t at which S buys and thus $\mu^s_{bh} < 1$ is unique. *Q.E.D.*

APPENDIX B: SUPPLEMENTARY MATERIAL FOR SECTION 6

PROOF OF LEMMA 5: (i) By standard results on MLRP and stochastic dominance, it must be that $\mathsf{E}[V|S_l] < \mathsf{E}[V|S_h]$. By a similar reasoning, at any history H^t, $\mathsf{E}[V|S_l, H^t] < \mathsf{E}[V|S_h, H^t]$ if the following MLRP condition holds at H^t: for any $S_l < S_h$ and any $V_l < V_h$,

$$
\text{(S4)} \qquad \frac{\mathsf{Pr}(S_h|V_h, H^t)}{\mathsf{Pr}(S_l|V_h, H^t)} > \frac{\mathsf{Pr}(S_h|V_l, H^t)}{\mathsf{Pr}(S_l|V_l, H^t)}.
$$

To show this, note first that $\mathsf{Pr}(V|H^t, S) = \mathsf{Pr}(V|S)\mathsf{Pr}(H^t|V)/\sum_{V'\in\mathbb{V}} \mathsf{Pr}(V'|S) \times \mathsf{Pr}(H^t|V')$. Then we have by the following manipulations that the MLRP condition $\frac{\Pr(S_h|V_h)}{\Pr(S_l|V_h)} > \frac{\Pr(S_h|V_l)}{\Pr(S_l|V_l)}$ implies the MLRP condition (S4) at any H^t:

$$
\begin{aligned}
&\Pr(S_l|V_l)\Pr(S_h|V_h) > \Pr(S_l|V_h)\Pr(S_h|V_l) \\
&\quad \Leftrightarrow \quad \Pr(V_l|S_l)\Pr(V_h|S_h) > \Pr(V_h|S_l)\Pr(V_l|S_h) \\
&\quad \Leftrightarrow \quad \frac{\Pr(V_l|S_l)\Pr(H^t|V_l)}{\sum_{\mathbb{V}}\Pr(V|S_l)\Pr(H^t|V)}\frac{\Pr(V_h|S_h)\Pr(H^t|V_h)}{\sum_{\mathbb{V}}\Pr(V|S_h)\Pr(H^t|V)} \\
&\qquad\qquad > \frac{\Pr(V_h|S_l)\Pr(H^t|V_h)}{\sum_{\mathbb{V}}\Pr(V|S_l)\Pr(H^t|V)}\frac{\Pr(V_l|S_h)\Pr(H^t|V_l)}{\sum_{\mathbb{V}}\Pr(V|S_h)\Pr(H^t|V)} \\
&\quad \Leftrightarrow \quad \Pr(V_l|H^t, S_l)\Pr(V_h|H^t, S_h) > \Pr(V_h|H^t, S_l)\Pr(V_l|H^t, S_h).
\end{aligned}
$$

(ii) Suppose contrary to the claim, that an informed trader with signal S_1 does not sell at some history H^t. Then by part (i), no informed trader sells at H^t. This implies that at history H^t, $\text{bid}^t = \mathsf{E}[V|H^t]$. But since, by part (i), $\mathsf{E}[V|H^t]$ exceeds $\mathsf{E}[V|S_1, H^t]$, we have $\text{bid}^t > \mathsf{E}[V|S_1, H^t]$. Hence, an informed trader with signal S_1 sells at H^t. This is a contradiction.

The proof that informed traders with signal S_3 always buy is analogous.

(iii) First we show that $\Pr(S_1|V_1) > \Pr(S_1|V_3)$. Suppose otherwise; thus $\Pr(S_1|V_1) \leq \Pr(S_1|V_3)$. Then the two MLRP conditions $\Pr(S_1|V_1)\Pr(S_2|V_3) > \Pr(S_2|V_1)\Pr(S_1|V_3)$ and $\Pr(S_1|V_1)\Pr(S_3|V_3) > \Pr(S_3|V_1)\Pr(S_1|V_3)$ imply, respectively, that $\Pr(S_2|V_1) < \Pr(S_2|V_3)$ and $\Pr(S_3|V_1) < \Pr(S_3|V_3)$. Hence, since $\Pr(S_1|V_1) \leq \Pr(S_1|V_3)$, we have $\sum_{i=1}^3 \Pr(S_i|V_3) > \sum_{i=1}^3 \Pr(S_i|V_1)$. But this contradicts $\sum_{i=1}^3 \Pr(S_i|V_j) = 1$ for every j.

The same argument can be applied to show that $\Pr(S_1|V_1) > \Pr(S_1|V_2)$ and $\Pr(S_1|V_2) > \Pr(S_1|V_3)$, and also in the reverse direction for $\Pr(S_3|V_1) < \Pr(S_3|V_2) < \Pr(S_3|V_3)$.

(iv) Consider any arbitrary history H^t and any two values $V_l < V_h$. By (ii), type S_1 always sells and type S_3 always buys. There are thus two cases for a buy at H^t: either only S_3 types buy or S_2 *and* S_3 types buy. In the former case, $\beta_i^t = \gamma + \mu\Pr(S_3|V_i)$. As S_3 is strictly increasing, there exits $\varepsilon > 0$ such that $\beta_h^t - \beta_l^t > \varepsilon$. In the latter case,

$$
\begin{aligned}
\beta_h^t - \beta_l^t &= \mu(\Pr(S_3|V_h) + \Pr(S_2|V_h) - \Pr(S_3|V_l) - \Pr(S_2|V_l)) \\
&= \mu\big(1 - \Pr(S_1|V_h) - (1 - \Pr(S_1|V_l))\big) \\
&= \mu(\Pr(S_1|V_l) - \Pr(S_1|V_h)).
\end{aligned}
$$

Since S_1 is strictly decreasing, there exists an $\varepsilon > 0$ such that $\beta_h^t - \beta_l^t > \varepsilon$.

By a similar reasoning it can be shown that there must exist $\varepsilon > 0$ so that $\sigma_l^t - \sigma_h^t > \varepsilon$. *Q.E.D.*

APPENDIX C: SUPPLEMENTARY MATERIAL FOR SECTION 8

PROPOSITION 6a: *Assume MLRP. Consider any finite history* $H^r = (a^1, \dots, a^{r-1})$ *at which the priors in the two markets coincide*: $q_i^r = q_{i,o}^r$ *for* $i = 1, 2, 3$. *Suppose that* H^r *is followed by* $s \geq 0$ *sales*; *denote this history by* $H^t = (a^1, \dots, a^{r+s-1})$. *If* $\sigma_1/\sigma_3 \geq \sigma_{1,o}/\sigma_{3,o}$, *then* $\mathsf{E}[V|H^t] < \mathsf{E}_o[V|H^t]$.

PROOF: First, note that, by (38) in the proof of Proposition 6, we have

$$\text{(S5)} \qquad \begin{aligned} &\sigma_3\sigma_{2,o} - \sigma_{3,o}\sigma_2 = -\mu^2\rho_{12}^{23} + \mu\gamma(\mathsf{Pr}(S|V_2) - \mathsf{Pr}(S|V_3)) < 0, \\ &\sigma_2\sigma_{1,o} - \sigma_{2,o}\sigma_1 > \sigma_3\sigma_{1,o} - \sigma_{3,o}\sigma_1. \end{aligned}$$

Also, since for herding, we require $\mathsf{E}[V|S, H^1] < \mathsf{bid}^1$, it follows from (37) that

$$\begin{aligned} &q_2^1 q_1^1[\sigma_2\sigma_{1,o} - \sigma_{2,o}\sigma_1] + q_3^1 q_2^1[\sigma_3\sigma_{2,o} - \sigma_{3,o}\sigma_2] \\ &\quad + 2q_3^1 q_1^1[\sigma_3\sigma_{1,o} - \sigma_{3,o}\sigma_1] > 0. \end{aligned}$$

But then by (S5), we have

$$\text{(S6)} \qquad \sigma_2\sigma_{1,o} - \sigma_{2,o}\sigma_1 > 0.$$

Since $\mathsf{E}[V|H^t] - \mathsf{E}_o[V|H^t]$ has the same sign as the expression in (43), by simple expansion of this expression we have that if $b = 0$, then $\mathsf{E}[V|H^t] - \mathsf{E}_o[V|H^t]$ has the same sign as

$$\begin{aligned} &q_2^r q_1^r \left\{ (\sigma_2\sigma_{1,o} - \sigma_{2,o}\sigma_1) \sum_{\tau=0}^{s-1} (\sigma_2\sigma_{1,o})^{s-1-\tau} (\sigma_{2,o}\sigma_1)^\tau \right\} \\ &\quad + q_3^r q_2^r \left\{ [(\sigma_3\sigma_{2,o}) - (\sigma_{3,o}\sigma_2)] \sum_{\tau=0}^{s-1} (\sigma_3\sigma_{2,o})^{s-1-\tau} (\sigma_{3,o}\sigma_2)^\tau \right\} \\ &\quad + 2q_3^r q_1^r \left\{ (\sigma_3\sigma_{1,o} - \sigma_{3,o}\sigma_1) \sum_{\tau=0}^{s-1} (\sigma_3\sigma_{1,o})^{s-1-\tau} (\sigma_3\sigma_1)^\tau \right\}. \end{aligned}$$

Rearranging, $\mathsf{E}[V|H^t] - \mathsf{E}_o[V|H^t]$ has the same sign as

$$\text{(S7)} \qquad \begin{aligned} &q_2^r q_1^r \frac{\sum_{\tau=0}^{s-1} (\sigma_2\sigma_{1,o})^{s-1-\tau} (\sigma_{2,o}\sigma_1)^\tau}{\sum_{\tau=0}^{s-1} (\sigma_3\sigma_{2,o})^{s-1-\tau} (\sigma_{3,o}\sigma_2)^\tau} [\sigma_2\sigma_{1,o} - \sigma_{2,o}\sigma_1] \\ &\quad + q_3^r q_1^r [\sigma_3\sigma_{2,o} - \sigma_{3,o}\sigma_2] \end{aligned}$$

$$+2q_3^r q_1^r \frac{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{1,o})^{s-1-\tau}(\sigma_3\sigma_1)^\tau}{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{2,o})^{s-1-\tau}(\sigma_{3,o}\sigma_2)^\tau}[\sigma_3\sigma_{1,o}-\sigma_{3,o}\sigma_1].$$

Further manipulations show that

$$\left(\frac{\sigma_1}{\sigma_3}\right)^s > \frac{\sum_{\tau=0}^{s-1}(\sigma_2\sigma_{1,o})^{s-1-\tau}(\sigma_{2,o}\sigma_1)^\tau}{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{2,o})^{s-1-\tau}(\sigma_{3,o}\sigma_2)^\tau}$$

$$\Leftrightarrow \quad \sum_{\tau=0}^{s-1}{\sigma_2}^{s-1-\tau}{\sigma_{2,o}}^\tau(\sigma_1\sigma_3)^\tau((\sigma_1\sigma_{3,o})^{s-1-\tau}-(\sigma_3\sigma_{1,o})^{s-1-\tau})>0.$$

Also, by assumption we have $\frac{\sigma_1}{\sigma_3} > \frac{\sigma_{1,o}}{\sigma_{3,o}}$. Therefore, we must have

$$\text{(S8)} \qquad \left(\frac{\sigma_1}{\sigma_3}\right)^s > \frac{\sum_{\tau=0}^{s-1}(\sigma_2\sigma_{1,o})^{s-1-\tau}(\sigma_{2,o}\sigma_1)^\tau}{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{2,o})^{s-1-\tau}(\sigma_{3,o}\sigma_2)^\tau}.$$

Similar manipulations show that

$$\left(\frac{\sigma_1}{\sigma_2}\right)^s < \frac{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{1,o})^{s-1-\tau}(\sigma_3\sigma_1)^\tau}{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{2,o})^{s-1-\tau}(\sigma_{3,o}\sigma_2)^\tau}$$

$$\Leftrightarrow \quad \sum_{\tau=0}^{s-1}{\sigma_3}^{s-1-\tau}{\sigma_{3,o}}^\tau(\sigma_1\sigma_2)^\tau((\sigma_2\sigma_{1,o})^{s-1-\tau}-(\sigma_1\sigma_{2,o})^{s-1-\tau})>0.$$

This together with (S6), implies that

$$\text{(S9)} \qquad \left(\frac{\sigma_1}{\sigma_2}\right)^s < \frac{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{1,o})^{s-1-\tau}(\sigma_3\sigma_1)^\tau}{\sum_{\tau=0}^{s-1}(\sigma_3\sigma_{2,o})^{s-1-\tau}(\sigma_{3,o}\sigma_2)^\tau}.$$

Also, since $\mathsf{E}[V|S, H^t] - \mathsf{bid}^t > 0$,

$$\text{(S10)} \qquad q_2^r q_1^r \left(\frac{\sigma_1}{\sigma_3}\right)^s [\sigma_2\sigma_{1,o} - \sigma_{2,o}\sigma_1] + q_3^r q_2^r [\sigma_3\sigma_{2,o} - \sigma_{3,o}\sigma_2] + 2q_3^r q_1^r \left(\frac{\sigma_1}{\sigma_2}\right)^s [\sigma_3\sigma_{1,o} - \sigma_{3,o}\sigma_1] < 0.$$

Then it follows from (S10), together with $\frac{\sigma_1}{\sigma_3} > \frac{\sigma_{1,o}}{\sigma_{3,o}}$, (S6), (S8), and (S9), that the expression in (S7) is negative. Thus $\mathsf{E}[V|H^t] - \mathsf{E}_o[V|H^t] < 0$ and the result follows. *Q.E.D.*

APPENDIX D: SUPPLEMENTARY MATERIAL FOR SECTION 9

We prove Lemma 7 and Theorem 3 of Section 9 in a more general setup than that described in the main body of the paper. This more general setup is of independent interest, as it allows for uncertainties other than those relating to the value of the asset.

Specifically, suppose that there are $N \geq 3$ states, where each state represents all exogenous variables that might influence the prices, and assume that there are N signals. Without any loss of generality, order the states such that $V_1 \leq V_2 \leq \cdots \leq V_N$, where V_j denotes the value of the asset in state $j = 1, \ldots, N$. Note that here, in contrast to the model in the text, we allow for the possibility that the asset has the same values in different states to reflect the idea that there may be factors, other than the value of the asset, that may influence prices. In particular, we assume that the asset can have at most $I \leq N$ different values. We denote the (public) probability of state j at date t by q_j^t and denote the likelihood of signal S in state j by $\mathsf{Pr}(S|j)$.

We also restrict ourselves to a symmetric structure with respect to the values and the initial beliefs on the distribution of values of the asset, as in the three state model of the paper. Formally, we assume that the values are distributed on a symmetrical grid; thus $V_j \in \{0, \mathcal{V}, \ldots, (I-1)\mathcal{V}\}$ for all $j = 1, \ldots, N$ and $\mathcal{V} > 0$. Further, for any $r = 1, \ldots, I$, let $C_r := \{j | V_j = (r-1)\mathcal{V}\}$ be the set of states with valuations $(r-1)\mathcal{V}$ and let $c_r := |C_r|$ be the number of states with valuation $(r-1)\mathcal{V}$. Assume (i) $q_j^1 = q_{N+1-j}^1$ for every $j \leq N/2$ and (ii) $c_r = c_{I+1-r}$ for every $r \leq I/2$.

We say that signal S is *negatively biased* if for all $j \leq \frac{N}{2}$, we have $\mathsf{Pr}(S|j) < \mathsf{Pr}(S|N+1-j)$.

Notice that when $c_r = 1$ for all r (hence $I = N$), this setup is identical to the one in the main text.

We first prove that any informed type buys initially if it has a negative bias and if there are enough noise traders.

LEMMA I: *Let S be negatively biased. Then $\mathsf{E}[V|S] < \mathsf{E}[V]$. Hence, there exists $\mu^i \in (0, 1]$ such that S sells at the initial history if $\mu < \mu^i$.*

PROOF: Without loss of generality, we present the proof only for the case when the number of value classes I is even so that $I = 2k$ for some integer k. Then by the symmetry of the prior, $\mathsf{E}[V] = \mathcal{V}(2k-1)/2$. Also, $\mathsf{E}[V|S] = \mathcal{V}\sum_{r=1}^{2k}(r-1)\,\mathsf{Pr}(r|S)$, where $\mathsf{Pr}(r|S) = \sum_{j\in C_r}\mathsf{Pr}(j|S)$. Thus, we need to show

$$\sum_{r=1}^{2k}(r-1)\,\mathsf{Pr}(r|S) < \frac{2k-1}{2}. \tag{S11}$$

Next, since (a) $\mathsf{Pr}(S|j) > \mathsf{Pr}(S|N+1-j)$, (b) $q_j^1 = q_{N+1-j}^1$, and (c) $c_r = c_{I+1-r}$, we have $\mathsf{Pr}(r|S) > \mathsf{Pr}(2k+1-r|S)$ for all $r < (2k+1)/2$. Using this and $\sum_{r=1}^{R}\mathsf{Pr}(r|S) = 1$, we have $\sum_{r=1}^{k}\mathsf{Pr}(r|S) > \frac{1}{2} > \sum_{r=k+1}^{2k}\mathsf{Pr}(r|S)$. Therefore,

$$(k-1) + \sum_{r=k+1}^{2k}\mathsf{Pr}(r|S) < (k-1) + \frac{1}{2} = \frac{2k-1}{2}. \tag{S12}$$

Then by (S11), it is sufficient to show that

$$\sum_{r=1}^{k}(r-1)\,\mathsf{Pr}(r|S) + \sum_{r=k+1}^{2k}(r-1)\,\mathsf{Pr}(r|S) < (k-1) + \sum_{r=k+1}^{2k}\mathsf{Pr}(r|S). \tag{S13}$$

But the second term on the left hand side of (S13) satisfies

$$\begin{aligned}
&\sum_{r=k+1}^{2k}(r-1)\,\mathsf{Pr}(r|S)\\
&\quad= \sum_{r=k+1}^{2k}\mathsf{Pr}(r|S) + \sum_{r=k+1}^{2k}(r-2)\,\mathsf{Pr}(r|S)\\
&\quad< \sum_{r=k+1}^{2k}\mathsf{Pr}(r|S) + (k-1)\,\mathsf{Pr}(k+1|S)\\
&\qquad + [(k-1)\,\mathsf{Pr}(k+2|S) + \mathsf{Pr}(k-1|S)]\\
&\qquad + [(k-1)\,\mathsf{Pr}(k+3|S) + 2\,\mathsf{Pr}(k-2|S)] + \cdots\\
&\qquad + [(k-1)\,\mathsf{Pr}(2k|S) + (k-1)\,\mathsf{Pr}(1|S)]\\
&\quad= \sum_{r=k+1}^{2k}\mathsf{Pr}(r|S) + (k-1)\sum_{r=k+1}^{2k}\mathsf{Pr}(r|S) + \sum_{r=1}^{k}(k-r)\,\mathsf{Pr}(r|S).
\end{aligned}$$

Therefore, the left hand side of (S13) is less than

$$\sum_{r=1}^{k}(r-1)\,\mathsf{Pr}(r|S)+\sum_{r=k+1}^{2k}\mathsf{Pr}(r|S)$$
$$+(k-1)\sum_{r=k+1}^{2k}\mathsf{Pr}(r|S)+\sum_{r=1}^{k}(k-r)\,\mathsf{Pr}(r|S)$$
$$=(k-1)+\sum_{r=k+1}^{2k}\mathsf{Pr}(r|S).$$

This demonstrates that (S13) holds. Hence we must have that $\mathsf{E}[V|S]<\mathsf{E}[V]$. To complete the proof of the lemma we also need to show that there exists $\mu^i\in(0,1]$ such that $\mathsf{E}[V|S]<\mathsf{bid}^1$ if $\mu<\mu^i$. As in Lemma 5, this follows immediately from $\mathsf{E}[V|S]<\mathsf{E}[V]$ and from $\lim_{\mu\to 0}\mathsf{E}[V]-\mathsf{bid}^1=0$. This completes the proof of Lemma I. *Q.E.D.*

Next, we turn to the switching of behavior. In our main characterization results for the three state–three signal case to obtain switching by a herding type, we assumed that the signal is more likely when the value is highest than when the asset has the middle value; for the switching by a contrarian type we assumed that the signal is more likely when the value is lowest than when the asset has the middle value. The analogues of those conditions to the current setting with N states and I liquidation values are

(S14) $\quad \mathsf{Pr}(S|j)>\mathsf{Pr}(S|i)\quad$ for all $j\in C_I,\ i\in C_{I-1}$,

(S15) $\quad \mathsf{Pr}(S|j)<\mathsf{Pr}(S|i)\quad$ for all $j\in C_1,\ i\in C_2$.

In the three state case, a negatively biased signal that satisfies (S14) is nU-shaped and a negatively biased signal that satisfies (S15) is nhill-shaped.

Next we show that if the probability of informed trading is sufficiently small, then conditions (S14) and (S15) can be used to establish the following lemma.

LEMMA II: (i) *Let S satisfy* (S14). *Then there exists* $\mu^s_{bh}\in(0,1]$ *such that* $\beta^t_i\,\mathsf{Pr}(S|j)-\beta^t_j\,\mathsf{Pr}(S|i)>0$ *for all* $i\in C_{I-1}$, $j\in C_I$, t, *and* H^t.

(ii) *Let S satisfy* (S15). *Then there exists* $\mu^s_{bc}\in(0,1]$ *such that* $\beta^t_i\,\mathsf{Pr}(S|j)-\beta^t_j\,\mathsf{Pr}(S|i)>0$ *for all* $i\in C_1$, $j\in C_2$, t, *and* H^t.

PROOF: We show (i); the argument for (ii) follows analogously. Fix any $j\in C_I$ and $i\in C_{I-1}$. For any date t and history H^t, let $\mathcal{S}^t$ be a set of signal types that buy at history H^t. Since

$$\beta_i^t \Pr(S|j) - \beta_j^t \Pr(S|i) = \Big(\gamma + \mu \sum_{S' \in \mathcal{S}^t} \Pr(S'|i)\Big) \Pr(S|j) - \Big(\gamma + \mu \sum_{S' \in \mathcal{S}^t} \Pr(S'|j)\Big) \Pr(S|i),$$

it follows that $\beta_i^t \Pr(S|j) - \beta_j^t \Pr(S|i) > 0$ is equivalent to

$$\Pr(S|j) - \Pr(S|i) > \frac{\mu}{\gamma}\Big(\sum_{S' \in \mathcal{S}^t} \Pr(S'|j) \Pr(S|i) - \sum_{S' \in \mathcal{S}^t} \Pr(S'|i) \Pr(S|j)\Big). \tag{S16}$$

By (S14), the left hand side of (S16) is positive. Also, since there is a finite number of signals, the expression in parentheses on right hand side of (S16) is uniformly bounded in t. Therefore, there must exist $\mu_{bh}^s \in (0, 1]$ such that for any $\mu < \mu_{bh}^s$, (S16) holds for all t and H^t. *Q.E.D.*

Next, we state our first characterization result in this general setup (it is equivalent to Lemma 7 when $I = N$).

LEMMA III: (i) *Suppose S is negatively biased and satisfies* (S14), *and let the following condition hold*:

(S17) $\forall \varepsilon > 0\ \exists H^t$ *such that* $q_i^t/q_j^t < \varepsilon$ *for all* $j \in C_{I-1} \cup C_I$ *and* $i \notin C_{I-1} \cup C_I$.

Then there exists a $\mu_{bh} \in (0, 1]$ *such that S buy herds if* $\mu < \mu_{bh}$.

(ii) *Suppose S is negatively biased and satisfies* (S15), *and let the following condition hold*:

(S18) $\forall \varepsilon > 0\ \exists H^t$ *such that* $q_i^t/q_j^t < \varepsilon$ *for all* $j \in C_1 \cup C_2$ *and* $i \notin C_1 \cup C_2$.

Then there exists a $\mu_{bc} \in (0, 1]$ *such that S is a buy contrarian if* $\mu < \mu_{bc}$.

PROOF: We show part (i); part (ii) follows analogously. Assume that $\mu < \mu_{bh} \equiv \min\{\mu_{bh}^i, \mu_{bh}^s\}$, where μ_{bh}^i and μ_{bh}^s are, respectively, the bounds on the size of the informed traders given in Lemmas I and II. Since S is negatively biased and $\mu < \mu_{bh}$, by Lemma I, S sells at the initial history.

Analogously to Lemma 9, by simple calculations, it can be shown that for any history H_t, $\mathsf{E}[V|S, H^t] - \mathsf{ask}^t$ has the same sign as

$$\sum_{i<j} (V_j - V_i) \frac{q_i^t q_j^t}{\rho_I^t \rho_{I-1}^t} (\beta_i^t \Pr(S|j) - \beta_i^t \Pr(S|i)), \tag{S19}$$

where $\rho_r = \sum_{j\in C_r} q_j^t$ is the probability that the valuation is $(r-1)\mathcal{V}$ for all r. Consider now all terms in (S19) that pertain to both C_I and C_{I-1}. These are

$$\text{(S20)} \qquad \sum_{j\in C_I}\sum_{i\in C_{I-1}} \frac{q_i^t q_j^t}{\rho_I^t \rho_{I-1}^t}(\beta_i^t \Pr(S|j) - \beta_j^t \Pr(S|i)).$$

Since $\mu < \mu_{bh}^s$, by Lemma II, there exists an $\eta > 0$ such that $\beta_i^t \Pr(S|j) - \beta_j^t \Pr(S|i) > \eta$ for all $i \in C_{I-1}$, $j \in C_I$. Thus

$$\text{(S21)} \qquad \text{(S20)} > \eta \cdot \sum_{j\in C_I}\sum_{i\in C_{I-1}} \frac{q_i^t q_j^t}{\rho_I^t \rho_{I-1}^t} = \eta.$$

Furthermore, note that $\mathsf{E}[V|H^t] > \rho_{I-1}^t \mathcal{V}(I-2) + \rho_I^t \mathcal{V}(I-1)$. Also, by the symmetries assumed, it must be that $\mathsf{E}[V] \le \mathcal{V}(I-2)$. Therefore, $\mathsf{E}[V|H^t] - \mathsf{E}[V] > \rho_I^t \mathcal{V} - (1-\rho_I^t - \rho_{I-1}^t)\mathcal{V}(I-2) = \mathcal{V}\sum_{j\in C_I} q_j - \mathcal{V}(I-2)\sum_{j\notin C_I \cup C_{I-1}} q_j$. This together with (S17) and finiteness of the state space imply that there exists a history H^t such that the following two conditions hold:

$$\text{(S22)} \qquad \mathsf{E}[V|H^t] > E[V],$$

$$\text{(S23)} \qquad \sum_{\substack{i<j,\\ i,j\notin C_{I-1}\cup C_I}} (V_j - V_i)\frac{q_i^t q_j^t}{\rho_I^t \rho_{I-1}^t}(\beta_i^t \Pr(S|j) - \beta_i^t \Pr(S|i)) > -\eta.$$

The latter, together with (S21), implies that at such a history (S19)> 0. Thus, by (S22), type S buy herds at H^t. This completes the proof of Lemma III. *Q.E.D.*

Notice that the above result is the analogue of Lemma 4 for our current setup with N and I values. Also, properties (S17) and (S18) are, respectively, analogous to (3) and (4) for our setup.

As with (3) and (4), conditions (S17) and (S18) are assumptions on endogenous variables. One restriction on the information structure that ensures these properties is MLRP. In particular, with MLRP, one can show (as in the three state case) that the probability of a buy is increasing in the liquidation values and the probability of a sale is decreasing in the liquidation values; these relationships in turn ensure (S17) and (S18).

LEMMA IV: *Suppose that the signals satisfy MLRP and assume that $S_1 < \cdots < S_N$. Then there exists $\delta < 1$ such that for all i,j with $i < j$ and all t, we have $\beta_i^t/\beta_j^t < \delta$ and $\sigma_j^t/\sigma_i^t < \delta$.*

PROOF: We will show only $\beta_1^t < \beta_2^t < \cdots < \beta_N^t$; the result for σ_i follows analogously. To show the former, observe that with MLRP signals, expectations are ordered in signals: $\mathsf{E}[V|H^t, S_i] > \mathsf{E}[V|H^t, S_j]$ if $i > j$. Thus, if signal type S_i buys, so will all $S_l > S_i$ and for any t and any $i < j$, $\beta_i^t - \beta_j^t$ has the same sign as

$$\begin{aligned}
&\sum_{l=k}^{N} \mathsf{Pr}(S_l|j) - \sum_{l=k}^{N} \mathsf{Pr}(S_l|i) \\
&\quad = 1 - \sum_{l=1}^{k-1} \mathsf{Pr}(S_l|i) - \left(1 - \sum_{l=1}^{k-1} \mathsf{Pr}(S_l|j)\right) \\
&\quad = \sum_{l=1}^{k-1} \mathsf{Pr}(S_l|i) - \sum_{l=1}^{k-1} \mathsf{Pr}(S_l|j) \quad \text{for some } k \leq N.
\end{aligned}$$

Since MLRP implies first order stochastic dominance and there are a finite number of signals it follows that there exists $\varepsilon > 0$ such that $\sum_{l=1}^{k-1} \mathsf{Pr}(S_l|i) - \sum_{l=1}^{k-1} \mathsf{Pr}(S_l|j) > \varepsilon$ for all k and $i < j$. But then $\beta_i^t / \beta_j^t < 1 - \varepsilon$. This completes the proof of Lemma IV. *Q.E.D.*

THEOREM 3a: *Assume that signals satisfy MLRP and let signal S be negatively biased.*

(a) *If* $\mathsf{Pr}(S|V_{N-1}) < \mathsf{Pr}(S|V_N)$, *then there exists* $\mu_{bh} \in (0, 1]$ *such that S buy herds if* $\mu < \mu_{bh}$.

(b) *If* $\mathsf{Pr}(S|V_1) < \mathsf{Pr}(S|V_2)$, *then there exists* $\mu_{bh} \in (0, 1]$ *such that S is a buy contrarian if* $\mu < \mu_{bh}$.

Note that Theorem 3a is more general than Theorem 3, as it applies to the general setup depicted in this Supplemental Material.

PROOF OF THEOREM 3a: (a) It remains to be shown that histories exists such that (S17) holds. Consider the infinite path consisting of only buys at every date. By MLRP and Lemma IV, there must exist $\delta \in (0, 1)$ such that for every H^t and for any i and j with $i < j$, we have $\beta_i^t / \beta_j^t < \delta$. Since $\frac{q_i^{t+1}}{q_k^{t+1}} = \frac{\beta_i^t}{\beta_k^t} \frac{q_i^t}{q_k^t}$, it then follows that q_i^t / q_j^t converges to zero along this infinite path of buys, for all $i \notin C_{I-1} \cup C_I$ and $j \in C_{I-1} \cup C_I$. This together with Lemma III(i) concludes the proof for the existence of buy herding.

Part (b) follows analogously. *Q.E.D.*

APPENDIX E: SUPPLEMENTARY MATERIAL FOR SECTION 10

In this section, we show how AZ's *composition uncertainty* can be accommodated within our N-state framework of the last section and why the types that herd in this setup also have U-shaped signals.

In AZ's setup, there are three liquidation values 0, 1/2, and 1, as in their basic example. When the liquidation values are 0 and 1, there are two levels of informativeness of the market W and P. Thus, there are five states $(0, W)$, $(0, P)$, $1/2$, $(1, P)$, and $(1, W)$, and we enumerate them by $1, 2, 3, 4$, and 5, respectively. Thus, in terms of the notation from the previous section of this Supplemental Material, $I = 3$, $\mathcal{V} = 1/2$, and $N = 5$. We also denote the states with valuation i by C_i. Therefore, $C_0 = \{(0, W), (0, P)\}$, $C_{1/2} = \{1/2\}$, and $C_1 = \{(1, W), (1, P)\}$.

In terms of the private information of the traders, AZ's description is as follows. There are two kinds of informed traders. All have a common partition of the liquidation values given by $\{(0, 1), 1/2\}$. When values 0 or 1 are realized, the two types have different precisions with respect to the two valuations. Specifically, high "quality" type h has precision p_h and low "quality" type l has precision p_l, with $1 \geq p_h > p_l > 1/2$. Thus, in this model there five signals. We denote the signal that confirms valuation $1/2$ by S_3, let S_1 and S_5 be the signals that high quality traders receive, and let S_2 and S_4 be the signals that low quality types receive. Finally, the proportion of different kinds of traders depends on the informativeness of the market; in particular, the likelihood of quality type $i = h, l$ occurring in market j is given by μ_i^j, $i \in \{h, l\}$ and $j \in \{W, P\}$, with $\mu_h^W + \mu_l^W = \mu_h^P + \mu_l^P = \mu$.

Using our notation, we can then describe the information structure as follows:

	$(0, W)$	$(0, P)$	$1/2$	$(1, P)$	$(1, W)$
$\mathsf{Pr}(S\|i)$	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
S_1	$\frac{\mu_h^W}{\mu} p_h$	$\frac{\mu_h^P}{\mu} p_h$	0	$\frac{\mu_h^P}{\mu}(1 - p_h)$	$\frac{\mu_h^W}{\mu}(1 - p_h)$
S_2	$\frac{\mu_l^W}{\mu} p_l$	$\frac{\mu_l^P}{\mu} p_l$	0	$\frac{\mu_l^P}{\mu}(1 - p_l)$	$\frac{\mu_l^W}{\mu}(1 - p_l)$
S_3	0	0	1	0	0
S_4	$\frac{\mu_l^W}{\mu}(1 - p_l)$	$\frac{\mu_l^P}{\mu}(1 - p_l)$	0	$\frac{\mu_l^P}{\mu} p_l$	$\frac{\mu_l^W}{\mu} p_l$
S_5	$\frac{\mu_h^W}{\mu}(1 - p_h)$	$\frac{\mu_h^W}{\mu}(1 - p_h)$	0	$\frac{\mu_h^W}{\mu} p_h$	$\frac{\mu_h^W}{\mu} p_h$

To demonstrate buy herding, consider signals S_1 and S_2 (sell herding arguments are analogous and involve considering S_4 and S_5). These signals are U-shaped in the sense that $\mathsf{Pr}(S_i|j) > \mathsf{Pr}(S_i|k)$ for every $i = 1, 2$, $j \in C_0 \cup C_1$, and $k \in C_{1/2}$. By appealing to the arguments of the previous section, the U-shaped nature of signals S_1 and S_2 allows us to establish buy herding as follows.

Take the case of S_1 and suppose that S_1 does not buy herd. Note also that S_1 is negatively biased (in the generalized sense defined in the last section, $\mathsf{Pr}(S_1|j) > \mathsf{Pr}(S_1|6 - j)$ for each $j = 1, 2$) and satisfies (S14). Therefore, if (S17) holds, Lemma III applies and types S_1 buy herd when there are a sufficient number of noise traders—a contradiction. The last step is to construct histories that satisfy (S17).

The proof of (S17) is as in Proposition 3. (Since S_3 is a hill-shaped signal, it corresponds to Case D2 of that proof.[1]) Specifically, one constructs a two-stage history. During the first stage, the actions are such that q_i^t/q_j^t for $i \in C_0$, $j \in C_{1/2}$ decrease while q_i^t/q_j^t, for $i \in C_0$, $j \in C_1$ do not change (during this stage, the actions correspond to those that S_3 will take). The second stage, involves buys only. During this stage, q_i^t/q_j^t for each $i \in C_0$, $j \in C_1$ decreases while q_i^t/q_j^t for $i \in C_0$, $j \in C_{1/2}$ may increase. Finally, the lengths of the two stages are chosen appropriately so that (S17) holds (if the second stage is long enough, then q_i^t/q_j^t for each $i \in C_0$, $j \in C_1$ is sufficiently small, and if the length of the first stage is sufficiently long relative to the second stage, then q_i^t/q_j^t for each $i \in C_0$, $j \in C_{1/2}$ is sufficiently small).

APPENDIX F: SUPPLEMENTARY MATERIAL FOR SECTION 11

F.1. *Simple History Dependence*

The order of trades and traders does not affect the price path as long as the model primitives do not allow any type of trader to change behavior. Clearly, herding or contrarian behavior involves such a change of behavior; changes from buying to holding or from selling to holding also qualify as a change of behavior.

Without changes in behavior, it suffices to study the order imbalance (number of buys minus number of sales) to determine prices, but with changes, the order of arrival matters a great deal. Consider the following numerical example[2] of an MLRP signal structure with an nU-shaped S_2:

Pr$(S\|V)$	V_1	V_2	V_3
S_1	$\frac{40}{49}$	$\frac{4}{49}$	0
S_2	$\frac{9}{49}$	$\frac{9}{490}$	$\frac{243}{12{,}250}$
S_3	0	$\frac{9}{10}$	$\frac{12{,}007}{12{,}250}$

$$\mu = \frac{1209}{1600}, \quad \mathbb{V} = (0, 10, 20), \quad \text{Pr}(V) = (1/6, 2/3, 1/6).$$

For illustrative purposes, assume that the first 15 traders are all informed and each signal S_i, $i = 1, 2, 3$, is received by 5 of the first 15 traders. Next, we compare the price paths for different arrival orders of these traders.

[1]This subcase proves the existence of the histories that yield (S17) for the case with two U-/hill-shaped signals with opposing biases and a hill-/U-shaped signal with a zero bias, and it is the case for which our results formally subsume AZ's example of event uncertainty.

[2]We chose the numbers so that there can be herding after a small number of trades.

Series 1—The arrival order is $5 \times S_1$–$5 \times S_2$–$5 \times S_3$ (meaning the first five receive S_1, the next five receive S_2, and the last five receive S_3). The S_1 types, who move first, all sell and thus the price drops. The S_2 types also sell and the S_3 types buy. Computations show that after these 15 trades the public expectation will drop from 10 to 0.15.

Series 2—$5 \times S_1$–$5 \times S_3$–$5 \times S_2$. Here the outcome is the same as in the previous series with S_1 traders selling, S_3 types' buying, and finally the S_2 types selling. The public expectation also drops from 10 to 0.15.

Series 3—$5 \times S_3$–$5 \times S_2$–$5 \times S_1$. The S_3 traders move first and buy. The S_2 types will now behave differently from the previous two series and will be buy herding. The public expectation now rises to about 13.5. Finally, the five S_1 type sell and then the public expectation drops to 10.31.

The difference between the outcome for Series 3 with those of Series 1 and 2 illustrates how the arrival order of traders matters: since there are S_2 types who trade, this type's change in trading mode (from selling to buying) strongly affects the price path.

Note, however, that even if there are no S_2 types directly involved in trading, the market maker has to consider the possibility that this type trades and thus has to account for this type's change of trading mode. To illustrate this, we next compare the outcome when the same number of buys and sales occurs, but in different orders.

Series 4—20 buys followed by 20 sales. After 20 buys, the public expectation is 15.36; after 20 subsequent sales, it is 3.12.

Series 5—20 sales followed by 20 buys. After 20 sales, the public expectation is 1.16×10^{-13}; after 20 subsequent buys, it is 10.0064.

In summary, the S_2 type can change trading modes in response to observing the order flow; thus the order flow affects prices and the frequency of different types of future trades. In the short run, the fluctuations may thus be influenced by the precise order of trades.

F.2. *Price Sensitivity*

To further elaborate on the price sensitivity induced by herding, we simulate price paths (Figure 1) using the following MLRP specification:

$$\text{(S24)}\quad \mu^s_{bh} = 0.7656 \equiv \mu_{bh}, \quad \mu^i = 0.9215, \quad \mathbb{V} = (0, 10, 20), \quad \mathsf{Pr}(V) = (1/10, 4/5, 1/10),$$

$\mathsf{Pr}(S\mid V)$	V_1	V_2	V_3
S_1	$\frac{40{,}049}{49{,}000}$	$\frac{4}{49}$	0
S_2	$\frac{8951}{49{,}000}$	$\frac{9}{490}$	$\frac{243}{12{,}250}$
S_3	0	$\frac{9}{10}$	$\frac{12{,}007}{12{,}250}$

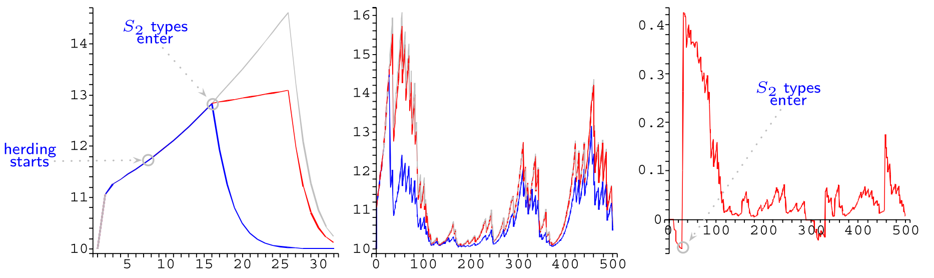


FIGURE 1.—Illustrations of the sensitivity in prices paths with and without herding.

In the left panel, there are two relevant price paths: the first (in gray) is for a setting with $\mu = \mu_{bh} - \varepsilon$, $\varepsilon = 1/10{,}000$; in other words, there is just enough noise so that herding is possible. The second price path (in red) is for $\mu = \mu_{bh}$ so that there cannot be herding.[3] The entry series for the graph is as follows: first, there is a long series of S_3 types, who all buy; this is followed by a group of S_2 types and eventually by some S_1 types. The point when S_2 types start entering is clearly marked; the S_1 types enter at the point when both curves peak. The point at which herding starts is marked too.

The series is constructed so that there are S_3 types who enter during herding. When the S_2 types enter, in the herding case, they buy; in the no-herding case, they hold. Even with holds, however, prices increase (this is due to the U-shaped c.s.d.).[4]

In the middle panel, we plot prices for the same specifications, this time for a random sequence of traders; both series have the same sequence of traders, but due to herding their actions may differ.[5] In the right panel, we plot the difference of the two rational price series from the middle panel. As the series with herding prices has more noise (because $\mu < \mu_{bh}$), initially, the price for the no-herding series is above the price of the herd series. Once herding starts (here after 8 trades) and once an S_2 type enters, this relation flips, illustrating that due to herding, prices move stronger in the direction of the herd than in the no-herding case.

F.3. *Does Herding Hamper Learning?*

To explore this issue, we use Monte Carlo simulations and compare the two scenarios outlined when discussing price sensitivity. That is, for the first series,

[3]The third price path (in blue) is for the case of the opaque economy as described in Section 8. For the opaque case, the differences in prices for the two levels of μ are negligible.

[4]The same simulation for the case of the opaque economy as described in Section 8 results in S_2 types selling and prices falling for both levels of μ.

[5]There is also a series for the opaque economy (Section 8) which, not surprisingly, is entirely below both rational series. Again, the opaque economy price series for $\mu = \mu_{bh}$ and $\mu = \mu_{bh} - \varepsilon$ are almost identical.

there is just enough noise so that buy herding can be triggered, $\mu = \mu_{bh} - \varepsilon$, $\varepsilon \approx 1/10{,}000$. In the second series, herding cannot occur, because there is too much informed trading, $\mu = \mu^{s}_{bh}$. We will refer to prices in the first setting as herding prices, irrespective of whether or not herding actually occurred; we refer to prices in the second setting as no-herding prices. Comparing the speeds of convergence for our two sets of simulations, we note the following two observations:

(i) If the true value is V_1 or V_2, then herding prices converge slower.

(ii) If the true value is V_3, then convergence with herding is faster.

These observations are based on the following: For the simulations, we again used the specification of the parameters given by (S24). Fixing the true liquidation values, we then drew 650 traders at random (noise and informed) assuming that $\mu_{bh} \approx 0.766$. Since the proportion of the informed agents μ is large—approximately three-quarters for both simulations—the 650 trades are almost always sufficient to obtain convergence to the true value. Next, we computed the time series of the transaction prices for both the herding and the no-herding case, and then recorded for each t and for both cases the absolute distance of the transaction price from the true value (which we know). We then repeated this procedure a large number of times, and calculated for each t and for each case the average distance from the true value. Since prices converge to the true value, these average distances decline in t. In the simulations, this distance declines approximately exponentially to zero. Thus the slope of the logarithm of the average distance measures the speed of convergence.

As the final step, we subtract at each t the log averages for the no-herding from the herding series. A positive number indicates that the herding series is slower, that is, that the average herding price is further away from the true value. Figure 2 plots these differences and the graphs are striking; they confirm our two observations mentioned above.[6]

To see the intuition for the these observations, compare the effects of buy herding on the herding and no-herding prices. First, when buy herding occurs, S_2 types buy in the herding case and thus there are more buys with herding than in the no-herding case. Second, in the case of a buy, prices in the herding case tend to be higher than in the no-herding case. Since the no-herding prices here are similar to or higher than those that arise in the opaque economy of Section 8 (only S_3 types buy in both cases), this second effect follows from the same reasoning used in the previous simulation to explain why, in the case of a buy, prices in the rational world, when herding starts, exceed those in the opaque economy (see Proposition 6.I(a)). Third, when there is a sale, prices in the herding and no-herding cases are almost identical and unaffected by buy herding. This is because in both cases only S_1 types sell: in the herding case,

[6]We have also made a formal analysis by regressing the log distance on time and, using the Chow test, checking whether one slope is steeper than the other. The results were highly significant.

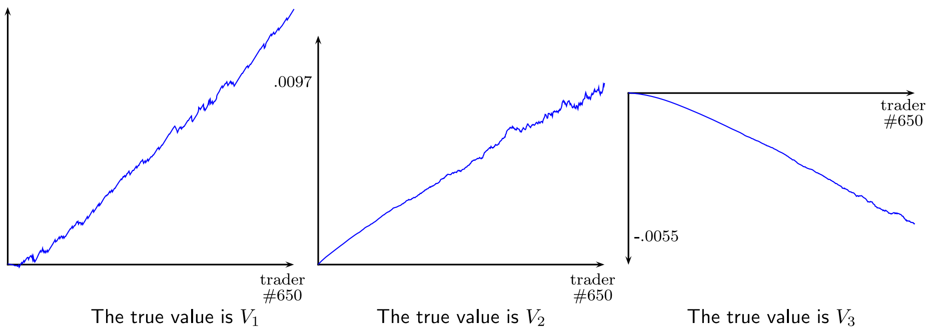


FIGURE 2.—The difference in speeds of convergence. Each graph plots the difference of the negative of the average log distance of the transaction prices of the herding and no-herding case. An up-sloping line thus indicates that for any t, herding prices are further from the true value than no-herding prices. All graphs are scaled to fit the page. The underlying signal distribution is listed in (S24).

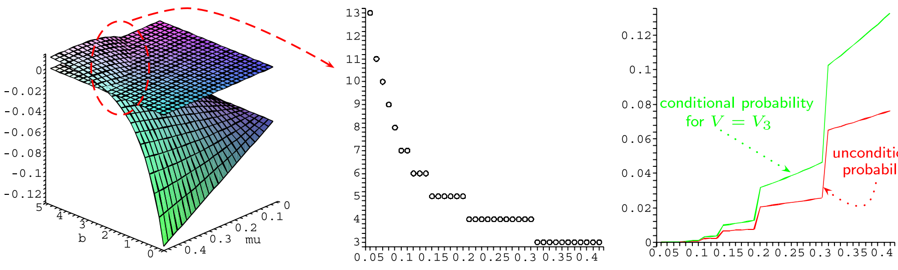
this is so by definition, and in the no-herding case, the S_2 type's expectation is almost equal to the ask price (expression (11) is almost zero) and thus larger than the bid price.[7]

Now it follows from the above that if the true value is V_1 or V_2, herding prices converge slower: during herding, herd buys move prices *away* from the true value by a larger magnitude and there are more such buys than in the no-herding case (sales have a similar effect in both cases). If, however, the true value is V_3, then once herding starts, prices in the herding case move up more strongly because of the first two effects and thus they move faster *toward* the true value. This leads to a higher speed of convergence in the herding case. Figure 2 documents these three cases.

F.4. *The Probability of the Fastest Herd*

The shortest sequence of trades that leads to buy herding is one with only buys; this is the "fastest" herd. We now want to get a sense of how likely this sequence is. Keeping the c.s.d. and the prior distribution fixed but varying the proportion of informed trading, we compute first how many buys are needed for buy herding to begin and then we determine how likely this sequence of buys is. The same type of analysis applies to sell herding.

[7]The herding and no-herding price paths may also differ even if no buy herding occurs (if S_2 types behave the same way in the two cases) because the proportions of informed trading μ are different for the two cases. In particular, when S_2 types do not buy herd, since μ is smaller in the herding case, each price movement in the herding price series is smaller than in the no-herding case, and as a result, speed of convergence is slower in the former series. However, since for the simulations, the difference between the values of μ is small ($\varepsilon = 1/10{,}000$), the consequence of this effect is small relative to the first two effects mentioned above.



Equation (11) Minimum b given μ Probability of b given μ

FIGURE 3.—Trades needed for fastest herd the probabilities for these trades. The left panel plots the value of expression (11), divided by $q_2 q_3$, as a function of μ, with $\mu \in (0, \mu_{bh})$, and of no-herd buys b. Whenever the bend curve crosses the 0 surface from below, herding is triggered. The middle panel computes the minimum integer number of no-herd buys that would trigger herding as a function of noise level μ. The right panel computes two probabilities: the first is the probability of having exactly the threshold number of buys at the beginning of trade (the thresholds are taken from the middle panel) conditional on the true state being V_3. The second probability is the unconditional likelihood of this threshold number. The plots in the right panel are functions of the μ. The signal distribution that underlies these plots is listed in (S24).

As was explained before, S_2 types buy at any history H_t if the expression in (11) is positive. As the amount of informed trading increases from 0 to μ_{bh}, there are then two opposing effects. First, as noise decreases, the positive term in expression (11) (the second term) becomes smaller. This implies that for any history, the difference between the market maker's and the S_2 type's expectation becomes smaller; thus to get buy herding, one needs more buys. Second, as noise decreases, the informational content of past behavior (public information) improves and this makes herding more likely. Formally, the first and third terms in (11), (the negative terms) decline as μ increases. This is because for any $i = 2, 3, \frac{\beta_1}{\beta_i} = \frac{\mu \Pr(S_3|V_1)+\gamma}{\mu \Pr(S_3|V_i)+\gamma}$ and $\frac{\partial(\beta_1/\beta_i)}{\partial\mu} = (\Pr(S_3|V_1) - \Pr(S_3|V_i))/\beta_i{}^2$ and thus, since S_1's c.s.d. is decreasing, $\frac{\partial(\beta_1/\beta i)}{\partial\mu} < 0$.

While we do not have an analytical result on the net effect of increasing μ from 0 to μ_b, in all numerical examples that we computed, the second effect dominates. Thus as noise trading declines (μ increases to μ_{bh}) it takes *fewer* buys to trigger buy herding. Figure 3 plots the minimum number of such consecutive time-zero buys needed to trigger buy herding for our simulations. As the amount of noise decreases, ex ante it gets more likely that these consecutive buy trades occur. (Figure 3's right panel illustrates these probabilities.)

Dept. of Economics, University of Toronto, Toronto, Ontario M5S3G7, Canada; andreas.park@utoronto.ca

and

Faculty of Economics, Cambridge University, Sidgwick Avenue, Cambridge CB3 9DD, United Kingdom; hamid.sabourian@econ.cam.ac.uk.

Manuscript received May, 2009; final revision received October, 2010.